Context-based Reciprocal Preferences and Peer Effects

Jose A. Carrasco[*] Rodrigo Harrison[†] Mauricio Villena[‡]

March 18, 2019

JEL codes: C72, A13
Keywords: Reciprocity, Endogenous Preferences, Peer Effects

Abstract

This paper develops a model that accounts for reciprocity and peers effects at the preference level to explain context dependent behavior. That is, it addresses the fact that people adjust the concern they express for others depending on who they interact with as well as on the strategic environment. We consider a strategic setup where heterogeneous players — intrinsic altruists, selfish and spiteful — are random ly matched in pairs. How much preferences adjust is our measure of peer effects. We show that peer effects exists — yet not necessarily in the same amount among all players — regardless of the information structure of the game and that their intensity reduces for extreme types; those too altruistic or too spiteful. Otherwise, for those players that are neither sufficiently altruistic nor sufficiently spiteful we might observe *preference-reversion* — the fact that due to large peer effects, an intrinsically altruistic (spiteful) player expects to behave spitefully (altruistically). Our model predicts that in frameworks characterized by positive externalities and strategic complements (substitutes) reciprocity choices become strategic substitutes (complements) and peer effects grow larger. A stochastically better opponents type distribution also leads to larger peer effects and more altruistic expected behavior. Weather peer effects are positive or negative crucially depends on how types compare to the type of strategic interaction of the game. When players types are not known, there is no preference-reversion and equilibrium preferences are as selfish as possible.

[*]Facultad de Ingeniería y Ciencias, Universidad Adolfo Ibañez, 2640 Diagonal Las Torres, Santiago, Chile, email: jose.carrasco@uai.cl, web: www.tonocarrasco.com

[†]Instituto de Economía, Pontificia Universidad Católica de Chile, Av. Vicuña Mackenna 4860, Macúl, Santiago, Chile, email: harrison@uc.cl, web: www.rodrigoharrison.com

[‡]Escuela de Negocios, Universidad Adolfo Ibañez, 2700 Diagonal Las Torres, Santiago, Chile, email: mauricio.villena@uai.cl, Web: www.mauriciovillena.com

1 Introduction

People adjust their *preferences and behavior* according to the *context* where they interact, and depending on whether they observe *others* doing so. As such, we observe people showing positive concern towards others in some occasions while behaving selfishly in others. Even more, despite the fact that cooperation and altruism lies at the heart of human lives, there might be some other contexts were we lash out to produce harm. This paper precisely pursues a theory to describe how people adjust their preferences — in particular the concern they show for others — depending on the specifics of the context. Our model is centered on the idea that preferences are subject to "peer effects", influenced by the behavior of others.[1] That is, as people interact and are constantly exposed to others, preferences are not immune to social influence and peers.

When people interact, the are two more salient elements that defines a context. First, the *information structure*. Typically, the concern for others depends on the information people count with about those that might benefit from their acts. For instance, individuals are more likely to increase their monetary donations when they know in how much need the recipient is. Second, the *strategic environment*. Our approach accounts for strategic interaction, and so preferences and peer effects are endogenous objects. We show that people behavior crucially depends on weather their actions are strategic complements or strategic substitutes.

In this paper, we argue that the *microstructure of preferences* is a key element that accounts for *context-dependent preferences* and peer effects that allows us to answer our main question: How do people adjust their behavior — in particular the concern they show for others — depending on the context? We present a model that accounts for pairwise random meetings (Okuno-Fujiwara and Postlewaite (1995)) between players with *interdependent preferences* that engage in a simultaneous move short run game (e.g. Cournot, Bertrand, etc.). Crucially, players also engage in long run strategic interaction at the preference level were peer effects arise: they play an underlying reciprocity game whose equilibrium determines preferences and concern towards others. That is, reciprocity is our model's key ingredient for endogenous preferences and peer effects.[2]

[1]Following Thöni and Gächter (2015), we do not distinguish between the terms "peer effects" and "social influence".

[2]The behavioral economics literature refer to a peer effect if an agent is influenced in his or her actions by what a comparison agent does, even if there are no material spillovers between agents and hence no direct preference links exist between peers. See Chen et al. (2010), Falk et al. (2013), Falk and Ichino (2006a), Frey and Meier (2004), Gächter and Thöni (2010a), Gächter et al. (2012), Gächter et al. (2013a), Ho and Su (2009),Kremer and Levy (2008), Mas and Moretti (2009), Mittone and Ploner (2011) and Sacerdote (2001).

More specifically, we assume that players inherit an *intrinsic preference or type* and that they *reciprocate* by choosing how much to weight his opponents type. As a result, players *induced preferences* are match-specific and might differ from their intrinsic ones; preferences now might be not only context-dependent, but also shaped by how peers reciprocate. The induced preferences summarize behavior and concern for others. To wit, how they differ from their intrinsic values is our main measure of peer effects; how a player preference might be influenced by the behavior and comparison with peers.

Absent a a specific microstructure for preferences means that the context only changes with the information structure. Regardless, in this case the prediction is unambiguous: there are no peer effects, and positive concern for others (altruism) only arises when the strategic context is one of strategic complements; otherwise, negative concern (spitefulness) arises. Instead, by solving our model, we find that there are peer effects and the fact that players preferences are altruistic, selfish or spiteful are context-dependent. They depend on players own intrinsic type, the information they count with about their opponents type — whether the opponents type is common knowledge or not — and the type of game players engage in — whether the short run game is one of strategic complements or strategic substitutes. This allows us to explain contextual behavior without assuming that intrinsic preferences change; what changes are the peers which in turn influence the preferences and behavior of the individual.

In Section 3 we fully characterize equilibrium reciprocity and induced preferences when the information structure is such that players types are common knowledge. To account for different strategic environments, we introduce a parameter $-1 \leqslant k \leqslant 1$, where $|k|$ captures the *degree of strategic interaction* in the short run game. We refer to the environment as one of *negative externalities and strategic substitutes* when $k < 0$ and one of *positive externalities and strategic complements* when $k > 0$. Regardless, we show that peer effects exists, yet not necessarily in the same amount among players (Proposition 1), and that this is true regardless of the information structure of the games.

In these cases players are able choose a type specific reciprocity strategy that is not only type dependent — on both, their own and their opponents intrinsic type — but that also depends on the strategic environment. We show that when the short run game is one of strategic substitutes with $k < 0$ (complements with $k > 0$), the reciprocity game becomes one of strategic complements (substitutes). Even more, as k rises, *the more intrinsically altruistic player grow less reciprocal, whereas the other grows more reciprocal.* That is, reciprocity best responses might be downward or upward sloping functions, depending on the strategic context. As a result, the fact that altruistic, spiteful of selfish preferences — and also their intensity — arise crucially depends on how players type compare to the degree of strategic interaction k.

We then measure peer effects by exploring how much preferences differ from their intrinsic value. In other words, do players expect to behave more altruistically or more spitefully than what they intrinsically are? Furthermore, is it possible for an intrinsically altruistic (spiteful) player to expect to behave spitefully (altruistically)? We compute the *expected interim preferences*, and we show that a stochastically better opponents type distribution — as well as larger values of k — leads to more altruistic expected behavior. It also depends on a players own type: both sufficiently altruistic or spiteful players expect to reduce the concern intensity for others. Even more, for those players that are neither sufficiently altruistic nor sufficiently spiteful we might observe *preference-reversion* — the fact that in expectation an intrinsically altruistic (spiteful) player behaves spitefully (altruistically). In these cases, induced preferences might reverse depending on the strategic context (Proposition 2).

In Section 4 we explore contexts that are better characterized with an information structure such that players types are unknown (i.e. the incomplete information case). We find that now the optimal reciprocity choice is a dominant strategy, and players reciprocate weighting their opponents expected type. Interestingly, and unlike the perfect information case, optimal reciprocity is not only independent of the opponent's type — as one might expected — but also of the strategic environment summarised in parameter k. In addition, contrary to the perfect information case, optimal reciprocity yields interim expected preferences that are *as selfish as they can be*. We find that *a necessary and sufficient condition for altruism (spite) to arise is either intrinsic altruism (spitefulness) on one player or that on average, players are altruists (spiteful)* (Proposition 3).

LITERATURE: Our framework is related to the literature on interdependent preferences as well as on the literature on peer effects. Unlike us, the interdependent preferences approach (Sobel (2005), Fehr and Schmidt (1999), Güth and Napel (2006), Charness and Rabin (2002), Koçkesen et al. (2000), Alger and Weibull (2013)) typically considers exogenously specified contexts or fixed preferences that are not influenced by others behavior. More recently Carrasco et al. (2018) explore the evolutionary stability of interdependent preferences in a context with perfect information and a strategic environment that shows negative externalities and strategic substitutes. We depart from these previous works in the sense that our framework is not evolutionary nor cultural transmission based, and it does consider players optimizing behavior and strategic interaction under a rich set of different contexts.

The literature on peer effects is large, yet has focused mostly on the field of education. The notion that peer effects are important to educational outcomes has been confirmed both theoretically (Arnott and Rowse (1987), Benabou (1993), Lazear (2001) and McMillan (2004)) and empirically (Sacerdote (2001)). Another branch of empirical literature on peer effects focused on other economically important settings also confirms that people's behavior is often shaped by what others do. Some examples include retirement savings decisions (Duflo and Saez (2002), Beshears et al. (2015)) corruption (Dong et al. (2012)), drugs and alcohol consumption (Kremer and Levy (2008), Gaviria and Raphael (2001)) and behavior at work (Bandiera et al. (2010), Guryan et al. (2009), Waldinger (2011), Ichino and Maggi (2000), Mas and Moretti (2009)).

Interdependent preferences and peer effects are well documented, and despite the fact they are both highly connected topics, there is still a large gap between them. In fact, the experimental work by Gächter et al. (2013b) suggests that social preferences models (instead of the social norm approach) provide a "parsimonious explanation" for peer effects. From the theoretical viewpoint, an exception is the recent work by Fershtman and Segal (2018), which we see as an attempt to connect preferences and social influence. They explore properties of the social influence functions and their effect on equilibrium behavior. Unlike us, they do not account for strategic behavior at the preference level. That is, even if players are aware of the fact that they influence others, they do not behave strategically, which one distinct property of our model.

In addition, our paper relates the experimental work that has tested peer effects (Zimmerman (2003), Falk and Ichino (2006b), Gächter and Thöni (2010b), Bougheas et al. (2013)). More recently, Thöni and Gächter (2015) present an experimental gift-exchange game with effort revision to study the role of peer effects in social preferences. They find that efforts are strategic complements, and that theories of reciprocity do not predict peer effects. Our model provides a tractable model model that analyzes peer effects and preferences in different strategic contexts: one of strategic complements or strategic substitutes. In another experiment on peer effects, Bardsley and Sausgruber (2005) finds that reciprocity accounts for roughly two thirds of the "crowding-in" tendency in public goods provision.

The paper is organized as follows. We present the model in Section 2 and offer theoretical predictions for reciprocity, preferences and peer effects in Section 3. We then explore peer effects when players types are not commonly known in Section 4. Finally, we present our conclusions in Section 5. All proofs are deferred to the appendix.

2 The Model

We consider two groups, indexed by $i, j \in \{1, 2\}$ with $i \neq j$, each with a continuum of players. Meetings happen continuously between players of different groups. Matched players independently choose $x_i \in \mathbb{R}_+$ and $x_j \in \mathbb{R}_+$ (e.g. prices, quantities, etc), respectively. They derive *material payoffs* $\pi_i(x_i, x_j) = x_i(1 - x_i + kx_j)$, where $|k|$ captures the *degree of strategic interaction*. However, preferences are *interdependent*, where β_{ij} summarizes player i concern over his opponents — player j — material payoff; that is, player i perceives *utility* $u_i(x_i, x_j) = \pi_i(x_i, x_j) + \beta_{ij}\pi_j(x_j, x_i)$. To guarantee $x_i, x_j \geqslant 0$ we let $-1 \leqslant k \leqslant 1$, that is the environment allows for both negative externalities and strategic substitutes ($k < 0$) or positive externalities and strategic complements ($k > 0$). This pairwise amount interaction defines a *short run normal form game* $G(\beta, k) = \{\{i, j\}, (x_i, x_j) \in \mathbb{R}^2_+, (u_i, u_j)\}$, whose equilibrium is described by the profile (x_i^*, x_j^*).

We aim to understand interdependent preferences microstructure and so our model accounts for two key ingredients. First, for player heterogeneity: within each group players vary by their *type* $\theta \sim F_i$ with continuous densities $f_i \equiv F_i'$ on $\Theta = [-1, 1]$ and with $\mathbb{E}_{\theta_i}[\theta_i] = \bar{\theta}_i \in \Theta$.[3] Each player learns his own type previous to the match. Second, for *peer effects* and *long run* strategic interaction at the preference level. We combine these ingredients adopting Levine (1998) specific functional form for preferences letting $\beta_{ij} \equiv \theta_i + \lambda_i(\theta_j - \theta_i)$, where $0 \leqslant \lambda_i \leqslant 1$ is a *reciprocity strategy* that weights players *intrinsic preferences* (i.e. types), an invariable component acquired through genetic inheritance. Since $\theta_i, \theta_j \in [-1, 1]$, preferences obey $\beta_{ij} \in [-1, 1]$. More specifically, given a type profile (θ_i, θ_j), preferences obey $\beta_{ij} \in [\min(\theta_i, \theta_j), \max(\theta_i, \theta_j)]$. We refer to player i as *intrinsically altruistic, selfish or spiteful* if $\theta_i > 0$, $\theta_i = 0$ or $\theta_i < 0$, respectively.

Equilibrium reciprocity — and induced preferences— arise as players exclusively pursue their *long run material payoff* $\pi_i(x_i^*, x_j^*) = \Pi_i$ by choosing their reciprocity coefficient λ_i. This pairwise interaction defines an underlying normal form *reciprocity (long run) game* $\Lambda(k) = \{\{i, j\}, (\lambda_i, \lambda_j) \in [0, 1]^2, (\Pi_i, \Pi_j)\}$, whose equilibrium determines how much preferences differ from their intrinsic values and so the size of the peer effects.

Later, in Section 4 we relax the information structure of the game and explore equilibrium reciprocity when types are not common knownledge.

[3]We use standard notation where $\mathbb{E}_Y[\cdot]$ denotes the expectation in the random variable Y.

3 Reciprocity under Complete Information

We first solve for equilibrium reciprocity when matched players types are common knowledge. That is, before the interaction each player not only knows his own type, but also his opponent's. We proceed in two steps; first solving the short run game $G(\beta, k)$ and then solving the reciprocity game $\Lambda(k)$, were peer effects arise. Crucially, in this case players are able choose a type specific reciprocity strategy $\lambda_i(\theta_i, \theta_j)$, and so they might reciprocate based not only on their own type, but also on their opponents . In other words, there might be peer effects if $\lambda_i(\theta_i, \theta_j) > 0$.

SHORT RUN GAME: Individuals care not only about their own material payoffs but also about the material payoffs of others.[4] In each pairwise meeting players maximize $u_i(x_i, x_j)$, strictly concave in x_i; then the FOC are necessary and sufficient for a maximum. Equilibrium strategies simultaneously solve each player optimization. Best responses are $x_i(x_j) = (1 + kx_j(1 + \beta_{ij}))/2$. As $k, \beta_{ij}, \beta_{ji} \in [-1, 1]$ then best responses have slopes in $[-1, 1]$. Two extreme cases arise: First, if $\beta_{ij} = \beta_{ji} = 1$ and $k = -1$, then $x_i^* + x_j^* = 1/2$, as best responses perfectly overlap. Second, if $k = \beta_{ij} = \beta_{ji} = 1$, both players best responses grow linearly without intersecting; precluding equilibrium existence. Otherwise, if $-1 < k < 1$ or $\min(\beta_{ij}, \beta_{ji}) < 1$, then the unique equilibrium is:

$$x_i^* = \frac{2 + k(1 + \beta_{ij})}{4 - k^2(1 + \beta_{ij})(1 + \beta_{ji})} \tag{1}$$

By standard derivation rules, we deduce that the amount x_i^* rises in β_{ji}. Intuitively, more concern by player j over player i's payoffs increases the latter's player strategy. In addition, we find that the difference $x_i^* - x_j^*$ is proportional to $k(\beta_{ij} - \beta_{ji})$; that is, if $G(\beta, k)$ is a game of strategic complements (substitutes), whoever exerts more concern towards his opponent's payoff will choose a larger (smaller) amount of x.

As $2 + k(1 + \beta_{ij}) > 0$, then (1) yields:[5]

$$\Pi_i \equiv \pi_i(x_i^*, x_j^*) = \frac{(2 + k(1 + \beta_{ij}))(2 + k(1 - \beta_{ij}(1 + k(1 + \beta_{ji}))))}{(4 - k^2(1 + \beta_{ij})(1 + \beta_{ji}))^2} > 0 \tag{2}$$

[4]Following Levine (1998), the linearity of the subjective utility in the opponents material payoff is a convenient approximation. Bester and Güth (1998), Bolle (2000), Possajennikov (2000) and Carrasco et al. (2018) have used the same setting to study the evolutionary stability of interdependent preferences.

[5]To show that $\Pi_i > 0$ we do: as $4 - k^2(1 + \beta_{ij})(1 + \beta_{ji}) > 0$ then $\Pi_i(\beta_{ij}, \beta_{ji}) > 0 \leftrightarrow 2 + k - k\beta_{ij}(1 + k(1+\beta_{ji})) > 0$. If $0 < k \leqslant 1$, as $-k(1+k(1+\beta_{ji})) < 0$ then $2+k-k\beta_{ij}(1+k(1+\beta_{ji})) > 2-(1+\beta_{ji})k^2 > 0$. If $-1 < k < 0$ then $-1 < 1 + k(1 + \beta_{ji}) \leqslant 1$ and as $\beta_{ij} \in [-1, 1]$ then $-1 \leqslant -\beta_{ij}(1 + k(1 + \beta_{ji})) \leqslant 1$ so $-k(1 - \beta_{ij}(1 + k(1 + \beta_{ji}))) \leqslant -2k < 2$.

Reciprocity Game: Exploiting (2), we see that absent a a specific microstructure for preferences, player i best response is $\beta_{ij} = (1+(2+k)k\beta_{ji})/(4+(2-k)k(1+\beta_{ji}))$, which yields a unique equilibrium $\beta^*_{ij} = \beta^*_{ji} = k/(2-k)$.[6]

Our microstructure adds a type dependent constraint to each player optimization problem — that restricts preferences in $\min(\theta_i, \theta_j) \leqslant \beta_{ij} \leqslant \max(\theta_i, \theta_j)$ — that might preclude the above symmetric equilibrium to arise. Clearly, the interesting cases require that $\theta_i \neq \theta_j$. To explore peer effects we compute the equilibrium reciprocity in these cases. That is, we solve the game $\Lambda(k)$. When players optimally choose how much reciprocity to exert, best responses are:[7]

$$\lambda_i(\lambda_j) = \frac{1}{(\theta_j - \theta_i)} \left(\frac{(1+\beta_{ji}(\lambda_j))\kappa(1+2\kappa)}{(1+\kappa)^2 + \kappa(1+\beta_{ji}(\lambda_j))} - \theta_i \right) \tag{3}$$

where $\kappa \equiv k/(2-k)$ is a *more convenient measure of strategic interaction.* Standard derivation rules yields $\partial\lambda_i(\lambda_j)/\partial\lambda_j > 0$ if $\kappa < 0$ and $\partial\lambda_i(\lambda_j)/\partial\lambda_j < 0$ if $\kappa > 0$. That is, *when the short run game $G(\beta, k)$ is one of strategic complements (substitutes), the reciprocity game $\Lambda(k)$ becomes one of strategic substitutes (complements).* Regardless, peer effects $\beta_{ij} - \theta_i$ might be positive or negative. We now compute equilibrium reciprocity. We omit the case $\kappa = -1/3$ ($k = -1$), as this was explored by Carrasco et al. (2018) in an evolutionary setting. Regardless, unlike them , induced preferences are symmetric despite the fact that players may *exert different reciprocity.*

Proposition 1 ***[Reciprocity Game** $\Lambda(k)$**]*** *For $\kappa \in (-1/3, 1]$: If $(\theta_i - \theta_j)(\theta_j - \kappa) \geqslant 0$ then $\lambda^*_i = 1$ and $\lambda^*_j = 0$ is the unique Nash equilibrium; otherwise, reciprocity $\lambda^*_i = (\theta_i - \kappa)/(\theta_i - \theta_j)$ and $\lambda^*_j = 1 - \lambda^*_i$ is the unique equilibrium. Additionally, for $\kappa = 1$, the equilibrium only exists if* $\max(\theta_i, \theta_j) < 1$.

Exploiting this result, we deduce that *reciprocity is not monotone neither in a players own type nor in the opponents type.* In other words, players do not grow more reciprocal the more intrinsically altruistic they are or their opponents are. In particular, it falls in θ_i when $\theta_i \leqslant \min(\kappa, \theta_j)$ and rises when $\theta_i \geqslant \max(\kappa, \theta_j)$; otherwise it equals zero, by Proposition 1. Equivalently, it is zero when $\theta_j < \min(\kappa, \theta_i)$ and jumps to one when $\min(\kappa, \theta_i) \leqslant \theta_j \leqslant \max(\kappa, \theta_i)$; otherwise falls in θ_j. As for the monotonicity in κ, we see that reciprocity λ_i is a piecewise linear function. As κ rises, *the more intrinsically altruistic player grow less reciprocal, whereas the other grows more reciprocal.* Regard-

[6]The second order condition for player i optimization is is $-k^2(4+(1+\beta_{ji})(2-k)k)^4/4(2+k(1+\bar{\beta}_{ji}))^2(2-k^2(1+\beta_{ji})) \leqslant 0$.

[7]This best response is derived by maximizing (2) in λ_i, subject to $0 \leqslant \lambda \leqslant 1$. This constraint means that $\beta_{ij} \in [\min(\theta_i, \theta_j), \max(\theta_i, \theta_j)]$.

less, in any equilibrium we have that $\lambda_i^* + \lambda_j^* = 1$, which guarantees that *ex-post peer effects exists yet, not necessarily among all players* and also *symmetric interdependent preferences.* To wit, players *exert different reciprocity, but behave the same.* Exploiting the inequalities statements of Proposition 1:

Corollary 1 *Induced interdependent preferences are:*

$$\beta_{ij}^*(\theta_i, \theta_j) = \beta_{ji}^*(\theta_j, \theta_i) = \min(\max(\kappa, \min(\theta_i, \theta_j)), \max(\theta_i, \theta_j))) \tag{4}$$

Preferences are piecewise linear and increasing functions of types and our more convenient measure of strategic interaction κ. Clearly, as induced preferences is a weighted average of types, when both players are intrinsically altruists (spiteful) then induced preferences are also altruist (spiteful). Otherwise, when players type significantly differ (only one player is an altruist or only one is spiteful), the type of game is relevant: *altruism only arises in games of strategic complements* (i.e $\beta_{ij}^* = \beta_{ji}^* > 0 \leftrightarrow \kappa > 0$). Even more, as κ rises in k we have that *the larger the degree of complementarity (substitutability), the more altruistic (spitefully) players preferences are.*

Regardless, the extreme cases of altruism or spitefulness where $\beta_{ij}^* = \beta_{ji}^* = 1$ or $\beta_{ij}^* = \beta_{ji}^* = -1$ are never induced interdependent preferences profiles. As a result, concern for others yields inefficient outcomes.[8] In fact, when $\kappa = k = 1$ there is no equilibrium in reciprocity, unless types can be further restricted. If $\max(\theta_i, \theta_j) = 1$, then it is optimal to choose $\lambda_i^* = 1$ for player i and $\lambda_j^* = 0$ for player j, if $\theta_i < \theta_j$, by Proposition 1. But then, $\beta_{ij}^*(\theta_i, \theta_j) = \beta_{ji}^*(\theta_j, \theta_i) = 1$ which yields no equilibrium in the short run game, by (1); a contradiction. To wit, players type need to be restricted to $\max(\theta_i, \theta_j) < 1$, in which case $\beta_{ij}^* = \max(\theta_i, \theta_j) = 1$ is the unique equilibrium.

When both players' types are sufficiently high (i.e. $\theta_i, \theta_j \geqslant \kappa$), the highest type player exerts what we call *strong reciprocity* ($\lambda^* = 1$) and the other is not reciprocal at all ($\lambda^* = 0$).[9] That is, *peer effects arise only in the higher type player.*[10] In this case, the induced preferences are $\beta_{ij}^*(\theta_i, \theta_j) = \beta_{ji}^*(\theta_j, \theta_i) = \min(\theta_i, \theta_j)$, by (4). That is, unlike spitefulness, altruism arises at their minimum possible intensity, at the lowest type player. As this behavior is more frequent the lower the value of κ, this is more likely in games of strategic substitutes. By the same token, when players' types are sufficiently

[8] The efficient outcome arises when $\beta_{ij}^* = \beta_{ji}^* = 1$ and so $x_i^* = x_j^* = 1/4(1-k)$ for $k \neq 1$ and $x_i^* + x_j^* = 1/2$ for $k = 1$.

[9] When $\kappa > 0$ this requires players to be sufficiently intrinsically altruistic; when $\kappa < 0$ this requires players to be not too intrinsically spiteful.

[10] In the education literature the define peer effects as :"For given educational resources provided to student A, if having student B as a classmate or schoolmate affects the educational outcome of A, then we regard this as a peer effect." (Chapter 20 in Epple and Romano (2011))

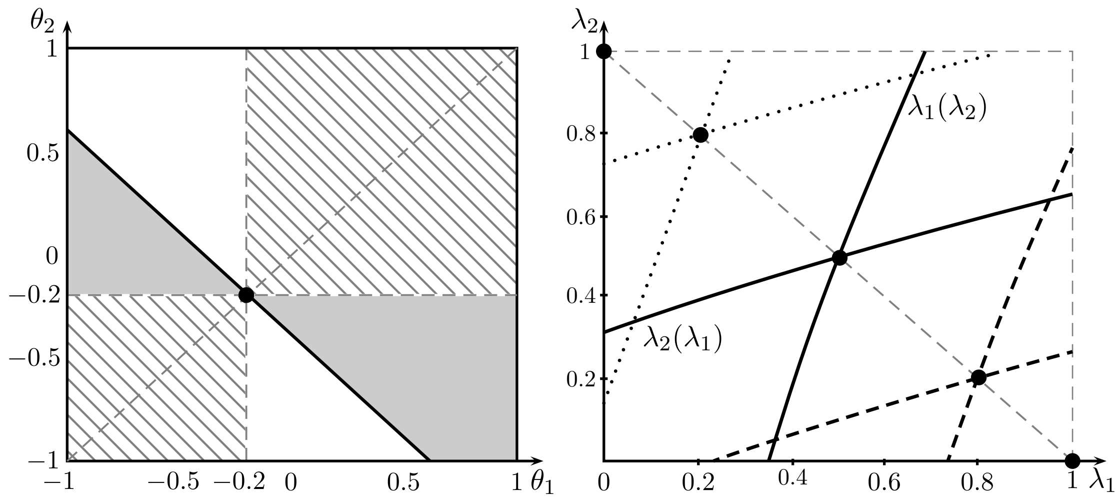

Figure 1: PREFERENCES AND RECIPROCITY FOR SUBSTITUTES ($\kappa = -1/5$). At left, the thick line is $\theta_1 + \theta_2 = -2/5$ in the players type space. In the upper dashed region behavior is symmetric at $\min(\theta_1, \theta_2)$, and at $\max(\theta_1, \theta_2)$ in the lower dashed region. In the remaining regions players behave spitefully at κ. In the gray region $\lambda_1^* > \lambda_2^*$ whereas in the white $\lambda_1^* < \lambda_2^*$. At right, the reciprocity are complements and the best responses slope upwards. Any equilibrium obeys $\lambda_1^* + \lambda_2^* = 1$. We depict equilibrium for (θ_1, θ_2) equal to $(0, -0.4)$ (solid), $(0, -0.25)$ (dashed) and $(-0.15, -0.4)$ (dotted).

low (i.e. $\theta_i, \theta_j \leqslant \kappa$), the lower type player exerts strong reciprocity and the other is not reciprocal. That is, *peer effects arise only in the lower type player.* Induced preferences are $\beta_{ij}^*(\theta_i, \theta_j) = \beta_{ji}^*(\theta_j, \theta_i) = \max(\theta_i, \theta_j)$, by (4). To wit, unlike spitefulness, altruism arises at the maximum possible intensity, at the highest players type. Otherwise, when $\min(\theta_i, \theta_j) < \kappa < \max(\theta_i, \theta_j)$, unlike $\kappa = 1$ case studied by Carrasco et al. (2018), best responses yield a unique solution with $0 < \lambda_i^*, \lambda_j^* < 1$ and $\beta_{ij}^*(\theta_i, \theta_j) = \beta_{ji}^*(\theta_j, \theta_i) = \kappa$. In this case, peer effects arise among all players.

In the *unique equilibrium* $\beta_{ij}^* = \beta_{ji}^* = \beta^*$ we obtain $\Pi_i = \Pi_j = (1 + \kappa)(1 + \kappa(1 - 2\beta^*))/4(1 - \kappa\beta^*)^2$, by (2), and so *altruism increases each player (and total) payoff.*[11] More specifically, when $-1/3 < \kappa < 1$, we have $x_i^* = x_j^* = 1/2(1-\kappa) > 0$ and $\Pi_i = \Pi_j = (1 + 2\kappa)/4(1 - \kappa^2) > 0$. In this case, *equilibrium outcomes do not vary symmetrically in* κ;[12] while $\beta_{ij}^* = \beta_{ji}^* \to 1$, $x_i^* = x_j^* \to \infty$ and $\pi_i^* = \pi_j^* \to \infty$ as $\kappa \to 1$, we have that $\beta_{ij}^* = \beta_{ji}^* \to -1/3$, $x_i^* = x_j^* \to 3/8$ and $\pi_i^* = \pi_j^* \to 3/32$ as $\kappa \to -1/3$. We have a symmetric equilibrium behavior, i.e $\lambda_i^* = \lambda_j^* = 1/2$ and $\beta_{ij}^* = \beta_{ji}^* = (\theta_i + \theta_j)/2$, only when $\theta_i + \theta_j = 2\kappa$, by Proposition 1.

We depict the last two paragraphs analysis in Figure 1 — when $G(\beta, k)$ is a game of strategic substitutes — and Figure 2 — when $G(\beta, k)$ is a game of strategic complements.

[11] As $\partial\Pi_i/\partial\beta^* = (1 - \beta^*)(1 + \kappa)\kappa^2/8(1 - \kappa\beta^*)^3 > 0$, this extends Proposition 1 in Bester and Güth (1998). Altruism increases efficiency, whereas spite reduces it.

[12] This is a common characteristic in games of strategic complementaries and substitutes (Milgrom and Roberts (1990), Frankel et al. (2003))

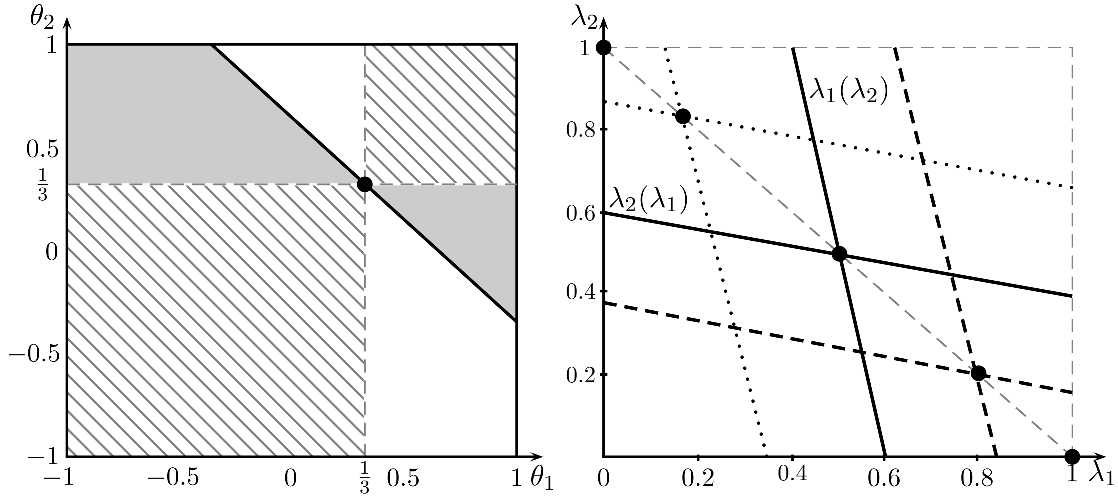

Figure 2: Preferences and Reciprocity for Complements ($\kappa = 1/3$). At left, the solid line is $\theta_1 + \theta_2 = 2/3$. In the upper (lower) dashed region the more (less) altruistic player exerts strong reciprocity; the other is not reciprocal. In the remaining parts behavior is symmetric and players behave altruistically at $\beta^*_{12} = \beta^*_{21} = \kappa$. In the gray region $\lambda^*_1 > \lambda^*_2$ whereas in the white $\lambda^*_1 < \lambda^*_2$. At right, the reciprocity are complements and the best responses slope downward. We depict the equilibrium for (θ_1, θ_2) equal to $(1/2, 1/6)$ (solid),$(-0.3, 0.5)$ (dashed) and $(0.4, 0)$ (dotted).

Interim Preference Distribution: Having solved our model, we are now able to further *quantify peer effects* by exploring how much interim preferences differ from their intrinsic value; that is, how much does $\bar{\beta}_i(\theta_i)$ differ from θ_i. As each player knows his own type previous to the match, we are able to characterise the *interim expected preferences* $\bar{\beta}_i(\theta_i) = \int_{-\infty}^{\infty} \beta^*_{ij}(\theta_i, \omega) dF_j(\omega)$ that summarized player i average concern on their opponent's payoffs.[13] Then, rewrite (4) to get:[14]

$$\bar{\beta}_i(\theta_i) = \begin{cases} \int_{-\infty}^{\infty} \min(\theta_i, \max(\kappa, \omega)) dF_j(\omega) & \text{for } \theta_i \geqslant \kappa \\ \int_{-\infty}^{\infty} \max(\theta_i, \min(\kappa, \omega)) dF_j(\omega) & \text{for } \theta_i \leqslant \kappa \end{cases} \tag{5}$$

The interim expected preference is increasing in θ_i; it is and convex for $\theta_i \leqslant \kappa$ and concave for $\theta_i \geqslant \kappa$. It is everywhere differentiable, expect (eventually) at $\theta_i = \kappa$.[15] It rises as F_j stochastically increases and rises in κ. That is *a stochastically better opponents type distribution leads to larger peer effects and more altruistic expected behavior.* For extreme types we have $\bar{\beta}_i(-1) = \mathbb{E}_{\theta_j}(\min(\kappa, \theta_j))$ and $\bar{\beta}_i(1) = \mathbb{E}_{\theta_j}(\max(\kappa, \theta_j))$, and so they vary differently as F_j grows more risky; while $\bar{\beta}_i(1)$ rises, $\bar{\beta}_i(-1)$ falls, by standard

[13] When $k = -1$, we restrict attention to the symmetric equilibrium, so that in all meetings preferences are symmetric and equal to $\beta^*_{ij}(\theta_i, \theta_j)$, as in (4).

[14] Write $\beta^*_{ij}(\theta_i, \theta_j) = \min(\max(\theta_i, \min(\kappa, \theta_j)), \max(\kappa, \theta_j)))$. As $\theta_i \geqslant \kappa \geqslant \min(\kappa, \theta_j)$, the first interval is obvious. For the second, observe that since $\theta_i \leqslant \kappa$ and $\min(\kappa, \theta_j) < \max(\kappa, \theta_j)$, then $\beta^*_{ij}(\theta_i, \theta_j) = \max(\theta_i, \min(\kappa, \theta_j))$.

[15] We have $\partial \bar{\beta}_i(\theta_i)/\partial \theta_i = F_j(\kappa)$ if $\theta_i < \kappa$ and $\partial \bar{\beta}_i(\theta_i)/\partial \theta_i = 1 - F_j(\kappa)$ if $\theta_i > \kappa$. To with, the derivative at $\theta_i = \kappa$ only exists when $1 - F_j(\kappa) = F_j(\kappa)$.

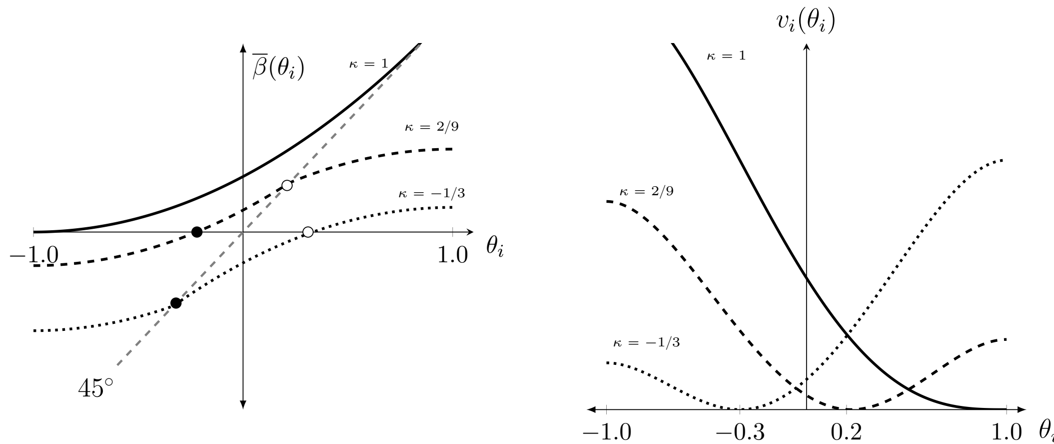

Figure 3: Interim Expected Preferences and its Variance. At left, the interim expected preferences. For $\theta_i \leqslant \underline{\theta}_i$ (black circles) then $\theta_i \leqslant \bar{\beta}_i(\theta_i) \leqslant 0$ and if $\theta_i \geqslant \bar{\theta}_i$ (white circles) then $0 \leqslant \bar{\beta}_i(\theta_i) \leqslant \theta_i$. Otherwise $\bar{\beta}_i(\theta_i) \leqslant \min(0, \theta_i)$ if $\kappa < 0$ and $\bar{\beta}_i(\theta_i) \geqslant \max(0, \theta_i)$ if $\kappa > 0$. At right, its variance. When $\theta_i = \kappa$, then $\beta^*_{ij} = \kappa$ and thus the variance is zero. We posit $\theta_i, \theta_j \sim u[-1, 1]$.

stochastic order logic. That is, spiteful (altruistic) players grow more spiteful (altruistic) as the type distribution is more dispersed.

Easily, use (5) to deduce that $\bar{\beta}_i(\theta_i) \geqslant \theta_i$ iff $\theta_i \leqslant \kappa$. That is, *there are positive peer effects (i.e. players exhibit more concern towards others) when their type exceeds our measure of strategic interaction κ; otherwise, peer effects are negative.*

Proposition 2 *For each $\kappa \in (-1/3, 1]$ there are values $\underline{\theta}_i(\kappa, F_j) = \underline{\theta}_i$ and $\bar{\theta}_i(\kappa, F_j) = \bar{\theta}_i$ obeying $-1 \leqslant \underline{\theta}_i < 0 < \bar{\theta}_i \leqslant 1$ such that: (a) if $\theta_i \geqslant \bar{\theta}_i$ then $0 \leqslant \bar{\beta}_i(\theta_i) \leqslant \theta_i$, (b) if $\theta_i \leqslant \underline{\theta}_i$ then $\theta_i \leqslant \bar{\beta}_i(\theta_i) \leqslant 0$ and (c) $\bar{\beta}_i(\theta_i) \geqslant \max(0, \theta_i)$ if $\kappa > 0$ and $\bar{\beta}_i(\theta_i) \leqslant \min(0, \theta_i)$ if $\kappa < 0$.*

We highlight three main conclusions, derived from Proposition 2. First, that *interim expected preferences* generically differ from the intrinsic types.[16] Second, that *peer effects intensity reduces for extreme types*: both *sufficiently altruistic or spiteful players reduce the concern intensity for others.* That is, a sufficiently altruistic player ($\theta_i \geqslant \bar{\theta}_i > 0$) expects to behave altruistically but not as much as he intrinsically are (i.e. peer effects are negative and $0 \leqslant \bar{\beta}_i(\theta_i) \leqslant \theta_i$), whereas a sufficiently spiteful player ($\theta_i \leqslant \underline{\theta}_i < 0$) expect to behave spitefully, but not as much as he are (i.e. peer effects are positive and $\theta_i \leqslant \bar{\beta}_i(\theta_i) \leqslant 0$). Third, for those players that are neither sufficiently altruistic nor sufficiently spiteful we might observe *preference-reversion* — the fact that due to large peer effects, an intrinsically altruistic (spiteful) player expects to behave spitefully (altruistically); formally, that $\theta_i\bar{\beta}_i(\theta_i) < 0$. In these cases (i.e. $\underline{\theta}_i < \theta_i < \bar{\theta}_i$), induced preferences might reverse depending on the strategic context: for $\kappa < 0$ ($\kappa > 0$) players

[16]They only coincide in the particular case when $\theta_i = \kappa$.

expect to behave spitefully (altruistically).

We compute the *induced preferences variance* $v_i(\theta_i) = \mathbb{E}_{\theta_j}(\beta_{ij}^{*2}) - \bar{\beta}_i(\theta_i)^2$ using (4):

$$v_i(\theta_i) = \begin{cases} 2\int_{\kappa}^{\theta_i}(\theta_i - \omega)F_j(\omega)d\omega - \left(\int_{\kappa}^{\theta_i} F_j(\omega)d\omega\right)^2 & \text{for } \theta_i \geqslant \kappa \\ -2\int_{\theta_i}^{\kappa}(\omega - \kappa)F_j(\omega)d\omega - \left(\int_{\theta_i}^{\kappa} F_j(\omega)d\omega\right)^2 & \text{for } \theta_i \leqslant \kappa \end{cases} \tag{6}$$

For $\theta_i < \kappa$, since $F_j(\omega) \leqslant 1$ and thus $\int_{\theta_i}^{\kappa} F_j(\omega)d\omega \leqslant \kappa - \theta_i$ we have $\partial v_i(\theta_i)/\partial\theta_i = 2F_j(\theta)\left(\theta_i - \kappa + \int_{\theta_i}^{\kappa} F_j(\omega)d\omega\right) < 0$. For $\kappa < \theta_i < 1$ we have $\partial v_i(\theta_i)/\partial\theta_i = 2(1 - F_j(\theta))\int_{\kappa}^{\theta_i} F_j(\omega)d\omega > 0$. At $\theta = \kappa$ then $v_i(\kappa) = 0$. That is, peer effects exhibit more variation when types differ more from κ. The variance depends not only on a player specific type, but also on the strategic context.

4 Reciprocity under Incomplete Information

Depending on the assumption about the size of the group and type of encounter, perhaps it is more accurate to modify the information structure of the game. We now assume that the specific type of the opponent is unknown and that each player only knows the distribution. Our model also captures these strategic environment, as long as we relax the complete information assumption. Aiming for this, we now assume that in each meeting, each player i only knows the opponents group distribution F_j. As a result, they cannot condition their behavior on who they meet. Reciprocity strategies will only depend on each player own type. In any case, we use Section 3 logic to solve for equilibrium behavior.

SHORT RUN GAME: Since in this game the opponent's type is unknown, players condition their preferences on their type. Specifically, let $b_i(\theta_i)$ be player i *incomplete information preference*. When meeting any other player, this preference will remain fixed at this stage. Accounting for our preference microstructure means that in each meeting, players choose how much reciprocity to exert by weighting their own type and the average opponents type; that is, $b_i(\theta_i) \equiv \theta_i + \lambda_i(\bar{\theta} - \theta_i)$. Then, player i ex post derived utility is $\pi_i(x_i, x_j) + b_i(\theta_i)\pi_j(x_j, x_i)$. To wit, we now face a *normal form Bayesian game*; each player only knows his type, and as such only his own interim preference. This pairwise interaction defines a *short run normal form game* $\mathcal{G}(b, k) = \{\{i, j\}, (x_i, x_j) \in \mathbb{R}_+^2, \Theta^2, \{F_i, F_j\}, (u_i, u_j)\}$, whose equilibrium is described by the profile $(x_i^*(\theta), x_j^*(\theta))$.

Let $x_i(\theta)$ be player i Bayesian equilibrium strategy. The *interim expected utility* is:

$$U_i(x_i|\theta_i) = x_i(1 - x_i + k\mathbb{E}_{\theta_j}[x_j(\theta_j)]) + b_i(\theta_i)\mathbb{E}_{\theta_j}[x_j(\theta_j)(1 - x_j(\theta_j) + kx_i)] \tag{7}$$

Let $\mathbb{E}_{\theta_j}[b_j(\theta_j)] = \bar{b}_j$. Then, solving for the equilibrium Bayesian strategies, we obtain:[17]

$$x_i^*(\theta_i) = \frac{(2 + k(b_i(\theta_i) - \bar{b}_i))(2 + k(1 + \bar{b}_j)) + 2k(\bar{b}_i - \bar{b}_j)}{2(4 - k^2(1 + \bar{b}_i)(1 + \bar{b}_j))} \tag{8}$$

Then, the expected long run profits, $\mathbb{E}_{\theta_j}[\Pi_i] = x_i^*(\theta_i)(1 - x_i^*(\theta_i) + k\mathbb{E}_{\theta_j}[x_j^*(\theta_j)])$, are:

$$\mathbb{E}_{\theta_j}[\Pi_i] = \frac{(4 + k(2 - k\bar{b}_i(1 + \bar{b}_j)))^2 - k^2(2 + k(1 + \bar{b}_j))^2(b_i(\theta_i))^2}{4(4 - k^2(1 + \bar{b}_i)(1 + \bar{b}_j))^2} \tag{9}$$

Reciprocity Game: Exploiting (9), we see that absent a preference microstructure, as $\mathbb{E}_{\theta_j}[\Pi_i]$ is a type-independent function, then $\bar{b}_i = b_i(\theta_i) \equiv b_i$ that are constant functions of the type. The question reduces now to find the value b_i that is optimal. For this, we set $\bar{b}_i = b_i$ and we write expected long run profits as:

$$\mathbb{E}_{\theta_j}[\Pi_i] = \frac{(2 + k(1 + b_i))(2 + k(1 - b_i(1 + k(1 + \bar{b}_j)))}{(4 - k^2(1 + b_i)(1 + \bar{b}_j))^2} \tag{10}$$

Interestingly, the expected profit function in (10) is analogous to (2), when player i preferences are $\beta_{ij} = b_i$ and player j's are $\beta_{ji} = \bar{b}_j$. As a result, the unique equilibrium is $b_i^* = b_j^* = \kappa$. Altogether, *absent a preference microstructure, either when preferences are known or not, predicted preferences are the same.* But, as in Section 3, our microstructure adds a type dependent constraint to each player optimization problem. Unlike in Section 3, now players weight their type and the average opponent's type, so $\min(\theta_i, \bar{\theta}_j) \leqslant b_i(\theta_i) \leqslant \max(\theta_i, \bar{\theta}_j)$.

We now solve our *reciprocity game* when types are not known. We redefine our game as $\Lambda(k) = \{\{i, j\}, (\lambda_i, \lambda_j) \in [0, 1]^2, \Theta^2, \{F_i, F_j\}, (\Pi_i, \Pi_j)\}$. For this, we optimize (9) in λ_i, taking as given $\bar{b}_j$. Exploiting our continuum type assumption, the specific choice of λ_i does not modify the expected preferences. To wit, we also consider $\bar{b}_i$ as a constant. As a result, we verify that conditional on his type θ_i, player i reciprocity choice is dominant. That is, players reciprocate independent of their opponents strategy λ_j. Analogous to the complete information case in Section 3, the interesting cases require that $\theta_i \neq \bar{\theta}_j$.

[17] For player i, the *best response* is $x_i(\theta_i) = (1 + k(1 + b_i(\theta_i))\mathbb{E}_{\theta_j}[x_j(\theta_j)])/2$. Taking expectations of player i best response, we obtain $\mathbb{E}_{\theta_i}[x_i(\theta_i)] = (1 + k(1 + \bar{b}_i)\mathbb{E}_{\theta_j}[x_j(\theta_j)])/2$. By the same logic, we obtain $\mathbb{E}_{\theta_j}[x_j(\theta_j)] = (1 + k(1 + \bar{b}_j)\mathbb{E}_{\theta_i}[x_i(\theta_i)])/2$ and so $\mathbb{E}_{\theta_j}[x_j(\theta_j)] = (2 + k(1 + \bar{b}_j))/(4 - k^2(1 + \bar{b}_i)(1 + \bar{b}_j))$. We obtain (8) by plugging $\mathbb{E}_{\theta_j}[x_j(\theta_j)]$ in player i best response.

By inspection of (9), optimal reciprocity is such that induced preferences $b_i(\theta_i)$ are as close as zero as possible; that is, as selfish as posible. We formalize these findings in Proposition 3 and Corollary 2:

Proposition 3 ***[Incomplete Information*** $\Lambda(k)$*] For $\theta_i \neq \bar{\theta}_j$ the dominant reciprocity strategy for each player i is: (a) $\lambda_i^* = \theta_i/(\theta_i - \bar{\theta}_j)$ if $\max(\theta_i, \bar{\theta}_j) \geqslant 0 \geqslant \min(\theta_i, \bar{\theta}_j)$, (b) $\lambda_i^* = 1$ if $\theta_i > \bar{\theta}_j \geqslant 0$ or $0 \geqslant \bar{\theta}_j > \theta_i$ and (c) $\lambda_i^* = 0$ if $0 \geqslant \theta_i > \bar{\theta}_j$ or $\bar{\theta}_j > \theta_i \geqslant 0$.*

Due to the linear preference microstructure, concavity of equation (9) in λ_i follows; to wit, the FOC is necessary and sufficient for a maximum. In this case we have that $\partial \mathbb{E}_{\theta_j}[\Pi_i]/\partial \lambda_i = 0 \leftrightarrow (\theta_i - \bar{\theta}_j) b_i(\theta_i) = 0$. Interestingly, optimal reciprocity is not only independent of the opponent's type — as one might expected — and dominant — as specified in Proposition 3 — but also of the strategic context summarised in parameter κ. As previously deduced, players interim preferences are as selfish as they can be:

Corollary 2 *The interim expected interdependent preferences are:*

$$b_i^*(\theta_i) = \min(\max(0, \min(\theta_i, \bar{\theta}_j)), \max(\theta_i, \bar{\theta}_j))) \tag{11}$$

For an intuition, observe that $k\mathbb{E}_{\theta_j}[x_j(\theta_j)]$ only varies with expected preferences $\bar{b}_i$ and $\bar{b}_j$, but not with $\lambda_i(\theta_i)$, by (8). As a result, players recognize that their reciprocity and induced preferences cannot influence their opponent's decision. Then, each player i maximizes $x_i(\theta_i)(A - x_i(\theta_i))$, where $A \equiv 1 + k\mathbb{E}_{\theta_j}[x_j(\theta_j)]$ acts as a *residual demand*, invariant to the reciprocity choice. It follows, by the well known monopoly rule, that $x_i(\theta_i) = A/2$ is optimal. Since player i best response dictates $x_i(\theta_i) = A/2 + k\mathbb{E}_{\theta_j}[x_j(\theta_j)] b_i(\theta_i)/2$, it is optimal to induce preferences that are as selfish as possible.

Relative to (4), we see that preferences in the incomplete information case are as if players were to match the average opponents type $\bar{\theta}_j$ when $\kappa \approx 0$. So, players behave as selfish as they can, given the constraints imposed by types. Easily, observe that $b_i^*(\theta_i) = 0$ whenever $\bar{\theta}_j \theta_i \leqslant 0$; otherwise, if $\max(\theta_i, \bar{\theta}_j) < 0$, induces $b_i^*(\theta_i) < 0$ and if $\min(\theta_i, \bar{\theta}_j) > 0$, induces $b_i^*(\theta_i) > 0$. That is, *a necessary and sufficient condition for altruism (spite) are both individual intrinsic altruism (spitefulness) and altruistic (spiteful) expected opponents type*. Otherwise, selfishness arises as induced preferences. This result is at odds with the findings of Ely and Yilankaya (2001) who for a general model of indirect evolution show that with incomplete information — when the preferences of the opponent are not known — only egoistic preferences (or preferences equivalent to them) survive evolution.

Peer effects are now measured by $b_i^*(\theta_i) - \theta_i$. Exploiting the above inequalities we deduce that the conditions for positive or negative peer effects are different than those in our complete information model. In particular, that *now there is no preference-reversion.* Assume $\theta_i \bar{\theta}_j \neq 0$. Then, we see that if $\theta_i < \min(0, \bar{\theta}_j)$ peer effects are positive and negative if $\theta_i > \max(0, \bar{\theta}_j)$. That is, a sufficiently spiteful player expect to behave spitefully, but not as much as he is (i.e. $0 > b_i^*(\theta_i) > \theta_i$), whereas a sufficiently altruistic player expects to behave altruistically but not as much as he intrinsically is (i.e. $\theta_i > b_i^*(\theta_i) > 0$). Otherwise — when $\min(0, \bar{\theta}_j) < \theta_i < \max(0, \bar{\theta}_j)$ — peer effects might be positive of negative depending on the sign of $\bar{\theta}_j$. Regardless, there is no preference-reversion. Whenever $\bar{\theta}_j \theta_i > 0$, then $\theta_i > b_i^*(\theta_i) > 0$ and so peer effects are negative. When $\bar{\theta}_j \theta_i < 0$, then $0 > b_i^*(\theta_i) > \theta_i$ and so peer effects are positive.

5 Conclusions

We formalize the notion that people adjust their preferences and behavior according to the context where they interact. At the heart of our theory is the idea that preferences are subject to reciprocity and peer effects, influenced by the behavior and comparison with others. Our approach also accounts for player heterogeneity, different strategic environments and information structures.

We measure peer effects by how much induced preferences differ from the their intrinsic values. We show that players reciprocity might differ leading to peer effects of different magnitude. This is regardless of the information structure of the game. We provide evidence that in some cases reciprocity will induce preferences such that due to large peer effects, an intrinsically altruistic (spiteful) player expects to behave spitefully (altruistically); that is, that we might observe preference-reversion. Weather peer effects are positive or negative crucially depends on how types compare to the type of strategic interaction of the game.

When there is incomplete information on other players types, we show that there is no preference-reversion and equilibrium preferences are as selfish as possible. Regardless, unlike Ely and Yilankaya (2001), altruistic as well as spiteful preferences might arise if we account for a microstructure of preferences and peer-effects.

Future extensions include replicating our analytical framework for more general matching technologies, other than pairwise random matching to examine the relationship between group size and peer effects.

6 Appendix

Lemma 1 *For $k \neq 0$, neither $\lambda_i = \lambda_j = 1$ nor $\lambda_i = \lambda_j = 0$ are equilibrium profiles.*

Proof: We show that no $\theta_i, \theta_j \in [-1, 1]$ solve simultaneously the FOC in (15), which are necessary for a maximum. In either case, the FOC dictate:

$$(\theta_i - \theta_j)(4\theta_j - k(1 + \theta_i)(2 - k)(b_3 - \theta_j)) \geqslant 0 \tag{12}$$

$$(\theta_j - \theta_i)(4\theta_i - k(1 + \theta_j)(2 - k)(b_3 - \theta_i)) \geqslant 0 \tag{13}$$

If $\theta_i > \theta_j$ then $k(1+\theta_i)(2-k)(b_3-\theta_j) \leqslant 4\theta_j < 4\theta_i \leqslant k(1+\theta_j)(2-k)(b_3-\theta_i)$ and so $k(2-k)(\theta_i - \theta_j)(b_3+1) < 0$; a contradiction for $k > 0$. The $\theta_j > \theta_i$ case is analogous. If $k < 0$ and $\theta_i > \theta_j$ then $\theta_i \in [0, b_3]$ clearly does not solve (13). Neither does $\theta_i \in (b_3, 1]$, as $k(1+\theta_j)(2-k)(b_3-\theta_i)$ rises linearly from 0 to $2k^2(1+\theta_j) < 4k^2 < 4$. The only candidates are $\theta_j < \theta_i < 0$. But in this case the FOC yield:

$$\frac{k(1+\theta_i)(2+k)}{4+k(1+\theta_i)(2-k)} \leqslant \theta_j \leqslant \frac{4\theta_i}{k(2-k)(b_3-\theta_i)} - 1 \tag{14}$$

This inequality limits are equal at $\theta_i = k/(2-k)$. Easily, their slopes are negative, and the upper limit slope exceeds the lower limit slope iff $-2 - 2\theta_i k + k^2 + \theta_i k^2 \geqslant 0$, which is iff $-k(2-k)(\theta_i + (2-k^2)/k(2-k)) \geqslant 0$. A contradiction. To wit, the interval in (14) is empty. The $\theta_j > \theta_i$ case is analogous. □

Proof of Proposition 1: Consider player $\theta_i \neq \theta_j$ maximization. As the Kuhn-Tucker FOC are necessary, we set up a Lagrangean $\mathcal{L} = \pi_i(x_i^*, x_j^*) + \gamma_0\lambda_i + \gamma_1(1-\lambda_i)$, where $\gamma_0, \gamma_1 \geqslant 0$ are the multipliers for $\lambda_i \geqslant 0$ and $\lambda_i \leqslant 1$. The FOC are:

$$\frac{(2+k(1+\beta_{ji}))((1+\beta_{ji})(2+k)k - \beta_{ij}(4+k(1+\beta_{ji})(2-k)))}{(4-k^2(1+\beta_{ij})(1+\beta_{ji}))^3} = \frac{(\gamma_1-\gamma_0)}{k^2(\theta_j-\theta_i)} \tag{15}$$

with $\gamma_0\lambda_i = 0$ and $\gamma_1(1-\lambda_i) = 0$. When $\lambda_i^* = 1$ and $\lambda_j^* = 0$ then $\gamma_0 = \gamma_1' = 0$, $\beta_{ij} = \beta_{ji} = \theta_j$ and $k^2(\theta_j - \theta_i)(k - \theta_j(2-k))/(2-k(1+\theta_j))^3(2+k(1+\theta_j)) = \gamma_1 = \gamma_0' \geqslant 0$, by (15). As $\theta_j \in [-1, 1]$ then $(2-k(1+\theta_j))(2+k(1+\theta_j)) > 0$ and so $(\theta_j - \theta_i)(k - \theta_j(2-k)) \geqslant 0$.

For uniqueness, we argue that only $\lambda_i^* = 1$, $\lambda_j^* = 0$ solves the Kuhn-Tucker conditions, and as maximum exists in $[0, 1]$, it is the unique maximum. We argue by contradiction, letting $(\theta_j - \theta_i)(k - \theta_j(2-k)) \geqslant 0$ and $\lambda_i^* < 1$ or $\lambda_j^* < 0$ or both.

Case 1: If $\lambda_i^* < 1$ and $\lambda_j^* > 0$: Then $\gamma_1 = \gamma_0' = 0$. If $\theta_i > \theta_j$ then $\theta_j \geqslant k/(2-k)$ and $\beta_{ij}, \beta_{ji} > k/(2-k)$. If $\theta_j > \theta_i$ then $\theta_j \leqslant k/(2-k)$ and $\beta_{ij}, \beta_{ji} < k/(2-k)$.

As $4 - k^2(1+\beta_{ij})(1+\beta_{ji}) > 0$ for $-1 < k < 1$, both players FOC in (15) yield:

$$(\theta_i - \theta_j)k(b_3 - \beta_{ij})(\beta_{ji} - b_1) \geqslant 0 \tag{16}$$

$$(\theta_i - \theta_j)k(b_4 - \beta_{ij})(\beta_{ji} - b_2) \geqslant 0 \tag{17}$$

with $b_1 = 4\beta_{ij}/(k(2+k-\beta_{ij}(2-k))) - 1$, $b_2 = (1+\beta_{ij})(2+k)k/(4+k(1+\beta_{ij})(2-k))$, $b_3 = (2+k)/(2-k)$, $b_4 = -4/k(2-k) - 1$ and:

$$\frac{(b_2 - b_1)(b_3 - \beta_{ij})(b_4 - \beta_{ij})}{(\beta_{ij}(2-k) - k)} = \frac{4(1+k)(2+k(1+\beta_{ij}))}{k^2(2-k)^2} \geqslant 0 \tag{18}$$

Observe that $b_3 > k/(2-k)$, $b_4 > b_3 \leftrightarrow k < 0$ and $b_4 > k/(2-k) \leftrightarrow k < 0$.

For $k < 0$ and $\theta_i > \theta_j$, then $b_4 > b_3 > k/(2-k)$ and $\beta_{ij}, \beta_{ji} > k/(2-k)$. If $b_3 > \beta_{ij} > k/(2-k)$ then (16), (17) and (18) yield $\beta_{ji} \leqslant b_1$. But as $\beta_{ij} = b_1$ at $\beta_{ij} = k/(2-k)$ and $\partial b_1/\partial \beta_{ij} = 4(2+k)/k(2+k-\beta_{ij}(2-k))^2 < 0$, then for $\beta_{ij} > k/(2-k)$ we have $\beta_{ji} < k/(2-k)$. A contradiction. If $b_4 > \beta_{ij} > b_3$ then (16), (17) and (18) yield $b_1 \leqslant \beta_{ji} \leqslant b_2$ and $b_2 \leqslant b_1$. A contradiction. If $\beta_{ij} > b_4$ then (16), (17) and (18) yield $\beta_{ji} \geqslant b_2$. But since $\beta_{ij} = b_2$ at $\beta_{ij} = k/(2-k)$ and $\partial b_2/\partial \beta_{ij} = 4(2+k)k/(4+k(1+\beta_{ij})(2-k))^2 < 0$, then for $\beta_{ij} > k/(2-k)$ we have $\beta_{ji} < k/(2-k)$. A contradiction.

If $k < 0$ and $\theta_j > \theta_i$, then $\beta_{ij}, \beta_{ji} < k/(2-k) < b_3 < b_4$. To wit (16), (17) and (18) dictate $\beta_{ji} \geqslant b_1$. But as $\beta_{ij} = b_1$ at $\beta_{ij} = k/(2-k)$ and $b_1 = \partial b_1/\partial \beta_{ij} < 0$, then $\beta_{ij} < k/(2-k)$ yield $\beta_{ji} > k/(2-k)$. A contradiction.

If $0 < k < 1$ then $b_3 > 1$ and $b_4 < -1$ so $b_2 \geqslant b_1$ iff $\beta_{ij} \leqslant k/(2-k)$. Now (16) and (17) dictate $(\theta_j - \theta_i)(\beta_{ji} - b_1) \leqslant 0$ and $(\theta_j - \theta_i)(\beta_{ji} - b_2) \geqslant 0$. When $\theta_i > \theta_j$ this reduces to $b_1 \leqslant \beta_{ji} \leqslant b_2$, and $b_1 \geqslant b_2$, by (18). A contradiction. Equivalently, if $\theta_j > \theta_i$ this reduces to $b_2 \leqslant \beta_{ji} \leqslant b_1$ and $b_2 \geqslant b_1$. A contradiction.

Case 2: If $\lambda_i^* < 1$ and $\lambda_j^* = 0$: Then $\gamma_1 = \gamma_1' = 0$, $\beta_{ji} = \theta_j$, $\theta_j < \beta_{ij} \leqslant \theta_i$ if $\theta_i > \theta_j$ and $\theta_i \leqslant \beta_{ij} < \theta_j$ if $\theta_j > \theta_i$. In this case the FOC yield (16) and the reversed inequality of (17). We now use the same logic of the previous case. For $k < 0$ and $\theta_i > \theta_j$, if $b_3 > \beta_{ij} > k/(2-k)$ then (16), (17) and (18) yield $b_2 \leqslant \beta_{ji} \leqslant b_1$ and $b_1 \leqslant b_2$. A contradiction. If $b_4 > \beta_{ij} > b_3$ then (16), (17) and (18) yield $\beta_{ji} \geqslant b_1$. But $\partial b_1/\partial \beta_{ij} < 0$ and $b_1 > 1$ at $\beta_{ij} = b_4$, so for $b_4 > \beta_{ij} > b_3$ we have $\beta_{ji} > 1$. A contradiction. If $\beta_{ij} > b_4$ then (16), (17) and (18) yield $b_1 \leqslant \beta_{ji} \leqslant b_2$ and $b_2 \geqslant b_1$. But since $\beta_{ij} = b_2$ at $\beta_{ij} = k/(2-k)$ and $\partial b_2/\partial \beta_{ij} < 0$, then $\beta_{ij} > k/(2-k)$ yields $\beta_{ji} < k/(2-k)$. A contradiction. For $k < 0$ and $\theta_j > \theta_i$, then $\beta_{ij}, \beta_{ji} < k/(2-k) < b_3 < b_4$. To wit (16), (17) and (18) dictate $b_1 \leqslant \beta_{ji} \leqslant b_2$ and $b_2 \leqslant b_1$. A contradiction.

If $0 < k < 1$ then $(\theta_j - \theta_i)(\beta_{ji} - b_1) \leqslant 0$ and $(\theta_j - \theta_i)(\beta_{ji} - b_2) \leqslant 0$ by (16) and (17) and $b_2 \leqslant b_1 \leftrightarrow \beta_{ij} \geqslant k/(2-k)$ by (18). When $\theta_i > \theta_j$ then $\beta_{ij} > k/(2-k)$, so the FOC reduce to $\beta_{ji} \geqslant b_1$. But $b_1 - \beta_{ij} = -(2+k(1+\beta_{ij}))(k-\beta_{ij}(2-k))/k(2+k-\beta_{ij}(2-k))$, then $\beta_{ij} \geqslant b_1 \leftrightarrow \beta_{ij} \leqslant k/(2-k)$. To wit, $\beta_{ij} < b_1 \leqslant \beta_{ji}$. A contradiction. Equivalently, if $\theta_j > \theta_i$ then $\beta_{ij} < k/(2-k)$ so the FOC yield $\beta_{ji} \leqslant b_1$. To wit, $\beta_{ji} \leqslant b_1 < \beta_{ij}$. A contradiction.

Case 3: If $\lambda_i^* = 1$ and $\lambda_j^* > 0$: Then $\gamma_0 = \gamma_0' = 0$, $\beta_{ij} = \theta_j$, $\theta_j < \beta_{ji} \leqslant \theta_i$ if $\theta_i > \theta_j$ and $\theta_i \leqslant \beta_{ji} < \theta_j$ if $\theta_j > \theta_i$. In this case the FOC yield the reversed inequality of (16) and (17). We now use the same logic of the previous case. For $k < 0$ and $\theta_i > \theta_j$, if $b_3 > \beta_{ij} > k/(2-k)$ then (16), (17) and (18) yield $b_1 \leqslant \beta_{ji} \leqslant b_2$ and $b_1 \leqslant b_2$. But since $b_2 = \beta_{ij}$ at $\beta_{ij} = k/(2-k)$ and $\partial b_2/\partial \beta_{ij} < 0$, then $\beta_{ij} > k/(2-k)$ yields $\beta_{ji} < k/(2-k)$. A contradiction. If $b_4 > \beta_{ij} > b_3$ then (16), (17) and (18) yield $b_2 \leqslant \beta_{ji} \leqslant b_1$ and $b_2 \leqslant b_1$. But as $b_1 = \beta_{ij}$ at $\beta_{ij} = k/(2-k)$ and $\partial b_1/\partial \beta_{ij} < 0$, then for $\beta_{ij} > k/(2-k)$ we have $\beta_{ji} < k/(2-k)$. A contradiction. If $\beta_{ij} > b_4$ then (16), (17) and (18) yield $b_2 \leqslant \beta_{ij} \leqslant b_1$ and $b_1 \geqslant b_2$. A contradiction.

If $k < 0$ and $\theta_j > \theta_i$, then $\beta_{ij}, \beta_{ji} < k/(2-k) < b_3 < b_4$. To wit (16), (17) and (18) dictate $b_2 \leqslant \beta_{ji} \leqslant b_1$ and $b_2 \leqslant b_1$. But as $b_1 = \beta_{ij}$ at $\beta_{ij} = k/(2-k)$ and $\partial b_2/\partial \beta_{ij} < 0$, then for $\beta_{ij} < k/(2-k)$ we have $\beta_{ji} > k/(2-k)$. A contradiction.

If $0 < k < 1$ then (16) and (17) dictate $(\theta_j - \theta_i)(\beta_{ji} - b_1) \geqslant 0$ and $(\theta_j - \theta_i)(\beta_{ji} - b_2) \geqslant 0$. When $\theta_i > \theta_j$ this reduces to $\beta_{ji} \leqslant b_2$, as $\beta_{ij} > k/(2-k)$. But $b_2 - \beta_{ij} = (2+k(1+\beta_{ij}))(k-\beta_{ij}(2-k))/(4+k(1+\beta_{ij})(2-k))$, then $\beta_{ij} \leqslant b_2 \leftrightarrow \beta_{ij} \leqslant k/(2-k)$. To wit, $\beta_{ji} \leqslant b_2 < \beta_{ij}$. A contradiction. If $\theta_j > \theta_i$ then $\beta_{ji} \geqslant b_2$. To wit, $\beta_{ij} < b_2 \leqslant \beta_{ji}$. A contradiction.

For the interior equilibrium, we intersect best responses in (3). This yields two candidates for equilibrium: $\lambda = (\theta_i - \kappa)/(\theta_i - \theta_j)$ and $\lambda' = (2 + \theta_i + 1/\kappa)/(\theta_i - \theta_j)$. We discard λ' as it is either negative or exceeds one. Letting $\lambda \in (0,1)$ yields $(\theta_i - \kappa)(\theta_j - \kappa) < 0$. Behavior is $\beta_{ij}^* = \beta_{ji}^* = \kappa$.□

Proof of Proposition 2: Integrating (5) by parts yields:

$$\bar{\beta}_i(\theta_i) = \begin{cases} \theta_i - \int_\kappa^{\theta_i} F_j(\omega)d\omega & \text{for } \theta_i \geqslant \kappa \\ \kappa - \int_{\theta_i}^{\kappa} F_j(\omega)d\omega & \text{for } \theta_i \leqslant \kappa \end{cases} \tag{19}$$

For $\kappa > 0$: as $\partial \bar{\beta}_i(\theta_i)/\partial \theta_i = 1 - F_j(\theta_i) \in [0,1)$ if $\theta_i \geqslant \kappa$, by (19), then $0 \leqslant \bar{\beta}_i(\theta_i) \leqslant \theta_i$. Hence, $\bar{\theta} = \kappa$. If $\theta_i \leqslant \kappa$ then $\bar{\beta}_i(\theta_i) \geqslant \theta_i$ and $\partial \bar{\beta}_i(\theta_i)/\partial \theta_i = F_j(\theta_i) \in [0,1)$. If $\bar{\beta}_i(-1) = \mathbb{E}(\min(\kappa, \Theta_j)) < 0$ then a unique $\underline{\theta}_i > -1$ solves $\bar{\beta}_i(\underline{\theta}_i) = 0$ and so $\theta_i \leqslant \bar{\beta}_i(\theta_i) \leqslant 0$ for $\theta_i \leqslant \underline{\theta}_i$. Easily, as $\bar{\beta}_i(0) > 0$, then $\underline{\theta}_i < 0$. If $\bar{\beta}_i(-1) \geqslant 0$ then $\underline{\theta}_i = -1$. Easily, $\bar{\beta}_i(\theta_i) \geqslant \max(0, \theta_i)$ if $\underline{\theta}_i \leqslant \theta_i \leqslant \bar{\theta}_i$.

For $\kappa < 0$: if $\theta_i \leqslant \kappa$ then $\theta_i \leqslant \bar{\beta}_i(\theta_i) \leqslant 0$ and so $\underline{\theta}_i = \kappa$, by (19). Next, for $\theta_i \geqslant \kappa$ we have $\partial\bar{\beta}_i(\theta_i)/\partial\theta_i = 1 - F_j(\theta_i) \in [0, 1)$ and $\bar{\beta}_i(0) < 0$, by (19). To wit, a unique $\bar{\theta}_i > 0$ solves $\bar{\beta}_i(\bar{\theta}_i) = 0$ and for $\theta_i \geqslant \bar{\theta}_i$ then $0 \leqslant \bar{\beta}_i(\theta_i) \leqslant \theta_i$. We have $\bar{\theta}_i < 1$ iff $\bar{\beta}_i(1) = \mathbb{E}(\max(\kappa, \Theta_j)) > 0$. Easily, $\bar{\beta}_i(\theta_i) \leqslant \min(0, \theta_i)$ if $\underline{\theta}_i \leqslant \theta_i \leqslant \bar{\theta}_i$.□

Proof of Proposition 3: We optimize (9) in λ_i. Observe that given our linear specification for preferences, $\partial b_i(\theta_i)/\partial\lambda_i = \partial\bar{\beta}_i^I(\theta_i)/\partial\lambda_i = \bar{\theta} - \theta_i$. Fix $\bar{b}_j$ and $\bar{b}_i$; then, computing the FOC we get:

$$\frac{\partial\mathbb{E}_j[\Pi_i]}{\partial\lambda_i} = \frac{-k^2(2 + k(1 + \bar{b}_j))^2\bar{\beta}_i^I(\theta_i)(\bar{\theta} - \theta_i)}{2(4 - k^2(1 + \bar{b}_i)(1 + \bar{b}_j))^2} = 0 \leftrightarrow \bar{\beta}_i^I(\theta_i) = 0$$

Clearly, if $\theta_i = \bar{\theta}$ then any λ_i is optimal. Otherwise, the solution is $\lambda_i^* = \theta_i/(\theta_i - \bar{\theta})$. To guarantee $0 \leqslant \lambda_i^* \leqslant 1$ we restrict types to $\max(\theta_i, \bar{\theta}) \geqslant 0 \geqslant \min(\theta_i, \bar{\theta})$. Otherwise, if $\theta_i > \bar{\theta} \geqslant 0$ or $0 \geqslant \bar{\theta} > \theta_i$ then $\lambda_i^* = 1$, and $\lambda_i^* = 0$ if $0 \geqslant \theta_i > \bar{\theta}$ or $\bar{\theta} > \theta_i \geqslant 0$. □

References

Alger, I. and J. W. Weibull (2013): "Homo moralis?preference evolution under incomplete information and assortative matching," *Econometrica*, 81, 2269–2302.

Arnott, R. and J. Rowse (1987): "Peer group effects and educational attainment," *Journal of Public Economics*, 32, 287–305.

Bandiera, O., I. Barankay, and I. Rasul (2010): "Social Incentives in the Workplace," *The Review of Economic Studies*, 77, 417–458.

Bardsley, N. and R. Sausgruber (2005): "Conformity and reciprocity in public good provision," *Journal of Economic Psychology*, 26, 664 – 681.

Benabou, R. (1993): "Workings of a city: location, education, and production," *The Quarterly Journal of Economics*, 108, 619–652.

Beshears, J., J. J. Choi, D. Laibson, B. C. Madrian, and K. L. Milkman (2015): "The effect of providing peer information on retirement savings decisions," *The Journal of finance*, 70, 1161–1201.

Bester, H. and W. Güth (1998): "Is altruism evolutionarily stable?" *Journal of Economic Behavior and Organization*, 34, 193 – 209.

Bolle, F. (2000): "Is altruism evolutionarily stable? And envy and malevolence?: Remarks on Bester and Güth," *Journal of Economic Behavior and Organization*, 42, 131–133.

Bougheas, S., J. Nieboer, and M. Sefton (2013): "Risk-taking in social settings: Group and peer effects," *Journal of Economic Behavior & Organization*, 92, 273 – 283.

Carrasco, J. A., R. Harrison, and M. Villena (2018): "Interdependent preferences and endogenous reciprocity," *Journal of Behavioral and Experimental Economics*, 76, 68 – 75.

Charness, G. and M. Rabin (2002): "Understanding Social Preferences with Simple Tests," *The Quarterly Journal of Economics*, 117, 817–869.

Chen, Y., F. M. Harper, J. Konstan, and S. X. Li (2010): "Social comparisons and contributions to online communities: A field experiment on movielens," *American Economic Review*, 100, 1358–98.

Dong, B., U. Dulleck, and B. Torgler (2012): "Conditional corruption," *Journal of Economic Psychology*, 33, 609 – 627.

Duflo, E. and E. Saez (2002): "Participation and investment decisions in a retirement plan: the influence of colleagues' choices," *Journal of Public Economics*, 85, 121 – 148.

Ely, J. C. and O. Yilankaya (2001): "Nash equilibrium and the evolution of preferences," *journal of Economic Theory*, 97, 255–272.

Epple, D. and R. E. Romano (2011): "Peer effects in education: A survey of the theory and evidence," in *Handbook of social economics*, Elsevier, vol. 1, 1053–1163.

Falk, A., U. Fischbacher, and S. Gächter (2013): "Living in two neighborhoods?social interaction effects in the laboratory," *Economic Inquiry*, 51, 563–578.

Falk, A. and A. Ichino (2006a): "Clean evidence on peer effects," *Journal of labor economics*, 24, 39–57.

——— (2006b): "Clean Evidence on Peer Effects," *Journal of Labor Economics*, 24, 39–57.

Fehr, E. and K. M. Schmidt (1999): "A Theory of Fairness, Competition, and Cooperation," *The Quarterly Journal of Economics*, 114, 817–868.

Fershtman, C. and U. Segal (2018): "Preferences and Social Influence," *American Economic Journal: Microeconomics*, 10, 124–42.

Frankel, D. M., S. Morris, and A. Pauzner (2003): "Equilibrium selection in global games with strategic complementarities," *Journal of Economic Theory*, 108, 1 – 44.

Frey, B. S. and S. Meier (2004): "Social comparisons and pro-social behavior: Testing" conditional cooperation" in a field experiment," *American Economic Review*, 94, 1717–1722.

Gächter, S., D. Nosenzo, and M. Sefton (2012): "The impact of social comparisons on reciprocity," *The Scandinavian Journal of Economics*, 114, 1346–1367.

——— (2013a): "Peer effects in pro-social behavior: Social norms or social preferences?" *Journal of the European Economic Association*, 11, 548–573.

——— (2013b): "Peer effects in pro-social behavior: Social norms or social preferences?" *Journal of the European Economic Association*, 11, 548–573.

Gächter, S. and C. Thöni (2010a): "Social comparison and performance: Experimental evidence on the fair wage–effort hypothesis," *Journal of Economic Behavior & Organization*, 76, 531–543.

——— (2010b): "Social comparison and performance: Experimental evidence on the fair wage–effort hypothesis," *Journal of Economic Behavior & Organization*, 76, 531–543.

Gaviria, A. and S. Raphael (2001): "School-Based Peer Effects and Juvenile Behavior," *The Review of Economics and Statistics*, 83, 257–268.

Guryan, J., K. Kroft, and M. J. Notowidigdo (2009): "Peer effects in the workplace: Evidence from random groupings in professional golf tournaments," *American Economic Journal: Applied Economics*, 1, 34–68.

Güth, W. and S. Napel (2006): "Inequality Aversion in a Variety of Games – An Indirect Evolutionary Analysis*," *The Economic Journal*, 116, 1037–1056.

Ho, T.-H. and X. Su (2009): "Peer-induced fairness in games," *American Economic Review*, 99, 2022–49.

Ichino, A. and G. Maggi (2000): "Work environment and individual background: Explaining regional shirking differentials in a large Italian firm," *The Quarterly Journal of Economics*, 115, 1057–1090.

Koçkesen, L., E. A. Ok, and R. Sethi (2000): "The Strategic Advantage of Negatively Interdependent Preferences," *Journal of Economic Theory*, 92, 274 – 299.

Kremer, M. and D. Levy (2008): "Peer effects and alcohol use among college students," *Journal of Economic perspectives*, 22, 189–206.

Lazear, E. P. (2001): "Educational production," *The Quarterly Journal of Economics*, 116, 777–803.

Levine, D. K. (1998): "Modeling Altruism and Spitefulness in Experiments," *Review of Economic Dynamics*, 1, 593 – 622.

Mas, A. and E. Moretti (2009): "Peers at work," *American Economic Review*, 99, 112–45.

McMillan, R. (2004): "Competition, incentives, and public school productivity," *Journal of Public Economics*, 88, 1871–1892.

Milgrom, P. and J. Roberts (1990): "Rationalizability, Learning, and Equilibrium in Games with Strategic Complementarities," *Econometrica*, 58, 1255–1277.

Mittone, L. and M. Ploner (2011): "Peer pressure, social spillovers, and reciprocity: an experimental analysis," *Experimental Economics*, 14, 203–222.

Okuno-Fujiwara, M. and A. Postlewaite (1995): "Social norms and random matching games," *Games and Economic behavior*, 9, 79–109.

Possajennikov, A. (2000): "On the evolutionary stability of altruistic and spiteful preferences," *Journal of Economic Behavior and Organization*, 42, 125 – 129.

Sacerdote, B. (2001): "Peer effects with random assignment: Results for Dartmouth roommates," *The Quarterly journal of economics*, 116, 681–704.

Sobel, J. (2005): "Interdependent preferences and reciprocity," *Journal of economic literature*, 43, 392–436.

Thöni, C. and S. Gächter (2015): "Peer effects and social preferences in voluntary cooperation: A theoretical and experimental analysis," *Journal of Economic Psychology*, 48, 72–88.

Waldinger, F. (2011): "Peer effects in science: Evidence from the dismissal of scientists in Nazi Germany," *The Review of Economic Studies*, 79, 838–861.

Zimmerman, D. J. (2003): "Peer Effects in Academic Outcomes: Evidence from a Natural Experiment," *The Review of Economics and Statistics*, 85, 9–23.